To appear in ApJ, CITA-1999-35
Preprint typeset using LaTeX style emulateapj v. 14/09/00

THE NATURE OF THE RADIATIVE HYDRODYNAMIC INSTABILITIES IN RADIATIVELY SUPPORTED THOMSON ATMOSPHERES

Nir J. Shaviv[1,2]
1. Theoretical Astrophysics 130-33, California Institute of Technology, Pasadena, CA 91125
2. Current Address: Canadian Institute for Theoretical Astrophysics, University of Toronto
60 St. George St., Toronto, ON M5S 3H8, Canada
shaviv@cita.utoronto.ca
To appear in ApJ, CITA-1999-35

ABSTRACT

Atmospheres having a significant radiative support are shown to be intrinsically unstable at luminosities above a critical fraction $\Gamma_{\rm crit} \approx 0.5-0.85$ of the Eddington limit, with the exact value depending on the boundary conditions. Two different types of absolute radiation-hydrodynamic instabilities of acoustic waves are found to take place even in the electron scattering dominated limit. Both instabilities grow over dynamical time scales and both operate on non radial modes. One is stationary and arises only after the effects of the boundary conditions are taken into account, while the second is a propagating wave and is insensitive to the boundary conditions. Although a significant wind can be generated by these instabilities even below the classical Eddington luminosity limit, quasi-stable configurations can exist beyond the Eddington limit due to the generally reduced effective opacity.

The study is done using a rigorous numerical linear analysis of a gray plane parallel atmosphere under the Eddington approximation. We also present more simplified analytical explanations.

Subject headings: Radiative transfer — hydrodynamics — instabilities — stars: atmospheres — stars: variables: other — stars: oscillations

1. INTRODUCTION

Luminous objects shining close to the Eddington limit are of special interest since they occur in a wide range of astrophysical contexts. As such, they have been a popular topic for observational and theoretical investigations. Luminosities in the range of the Eddington limit appear in the very luminous stars populating the top of the HR diagram, in novae during the critical phases of the eruption and also in accreting objects of both galactic and extra-galactic scales. These very luminous objects usually display a great diversity of phenomena ranging from erratic variability to spatial structures and winds. Consequently, the theoretical modeling requires the incorporation of hydrodynamics, radiation and sometimes even magnetic fields. Undoubtedly, much of the interesting behavior stems from the large, dynamically important luminosity of these objects – luminosities that are close to the Eddington limit.

The classical Eddington limit is the upper limit possible to the luminosity of an object in hydrostatic balance (Eddington 1926). It is the luminosity for which the radiative force upwards balances the gravitational force downwards assuming the opacity of the (fully ionized) gas attains its minimal possible value, that is, the Thomson scattering opacity for non relativistic electrons. For a homogeneous static configuration, it is the upper luminosity limit because above it there is no consistent solution to the hydrostatic equations. An atmosphere that violates this upper limit will then simply blow itself apart – an optically very thick wind is then unavoidable and the object will evaporate on a relatively short time scale. When the opacity is larger than the Thomson scattering one, as is often the case, one can define a modified Eddington limit that is smaller than the 'classical' value (e.g., Humphreys & Davidson 1984, Lamers 1986, and Appenzeller 1986). The modified Eddington luminosity was used to explain the upper luminosity limit observed for stars and the variability that these stars exhibit. The meaning of the modified Eddington luminosity should however be taken cautiously, because as one approaches the modified Eddington limit, the increased scale height and reduced density imply that in any fully ionized atmosphere, the opacity law is reduced to that of Thomson scattering (Humphreys & Davidson 1994). Namely, as the modified limit is approached, it itself approaches the classical limit such that the modified Eddington limit is probably not an absolute luminosity limit.

On top of the interesting questions on the nature of atmospheres when their luminosity approaches the Eddington luminosity, exists the fundamental question: 'Is the Eddington luminosity the maximal luminosity under all conditions?'. Part of the answer was already answered in the negative by Shaviv (1998). Here we elaborate and extend the analysis and demonstrate that the classical Eddington limit is the limit under a very restrictive set of assumptions. Still we would like to emphasize that a significant factor in turning the Eddington limit into a major tool in setting astrophysical limits is the simplicity of the classical expression, namely $\mathcal{L}_E = 4\pi cGM/\sigma_{\rm Thomson}$.

There are several reasons why the study of radiative instabilities is important. Most, if not all of the objects shining close to the Eddington limit show variability. Novae for example, not only show very high (occasionally super-Eddington) luminosities over extended periods of time, they also show clumpy ejected material (e.g., the high resolution HST images, Shara et al. 1997) which could be explained using radiatively driven instabilities. Another example are the super-Giants populating the top of the HR diagram. Quite a few stars with $M \sim 100 M_\odot$ and

$L \sim 10^6 L_\odot$ exist across a wide range of spectral types and exhibit various degrees of variability. For example, several of the numerous O3 type stars at 30 Doradus are so bright that their luminosity is clearly very close to their Eddington limit (e.g., Massey & Hunter 1998). Less blue (and less young) are the Luminous Blue Variables (LBVs) termed so because of their often eruptive behavior in which their luminosity and rate of mass ejection can increase while their effective temperature decreases (see for example the review by Humphreys & Davidson 1994). When these objects are in their eruptive state, at least some of them clearly surpass the Eddington limit for significant lengths of time. For example, η Car in its eruption between 1830 and 1860 released roughly 10^{49} $ergs$ of light over a period much longer than its dynamical time scale (e.g., van Genderen & Thé 1985), yielding an average luminosity that is more than an order of magnitude higher than the Eddington limit (provided that the mass of η Car is of the order of $100 M_\odot$). The variability of these objects which is probably a consequence of the high luminosity, is therefore interesting to understand.

An additional type of interesting objects are 'very' and 'super massive objects' (VMO's and SMO's) that might have existed as massive pop III stars in the young universe. These massive stars are supposed to shine at luminosities very near the Eddington limit with the more massive of them at virtually the Eddington limit itself. Since the metalicity of these objects is extremely small, the analysis carried out here, which assumes that the opacity is dominated by Thomson scattering, is particularly relevant to these objects.

Super-critical accretion disks are an altogether different type of relevant luminous objects. Irrespective on whether they advect most of the excess accreted mass and energy above the Eddington rate into the black hole or whether the excess mass is blown away, each element of these disks should shine very close to local effective Eddington limit as these disks are always radiation pressure dominated.

Various works have shown that atmospheres shining close to the Eddington limit exhibit a wide range of instabilities as well as other phenomena such as the generation of strong winds. For example, a large radiation pressure contribution reduces the adiabatic index. The relatively 'loose' structure becomes unstable against convection when the adiabatic index decreases due to ionization (e.g., Ledoux 1965, Stothers & Chin 1993, and Wagenhuber & Weiss 1994).

Joss, Salpeter & Ostriker (1973) analyzed the appearance of convection at high stellar luminosities assuming that the condition for convective instability is the classical Schwarzchild condition, namely the temperature gradient must be higher than the adiabatic one. The main concern of Joss et al. was the energy flux and its relation to the Eddington luminosity. Here we are predominately interested in the *conditions* for the stability under high radiative fluxes. The main result of Joss et al. was that as the energy flux through an atmosphere is increased, convection arises before the Eddington limit is reached. Moreover, the convective flux will always grow enough such that the remaining radiative flux is less than Eddington, *as long as convection is efficient*. Namely, the interior of stars which are dense, will never witness super Eddington fluxes. It is only near the surface where super Eddington fluxes can be reached. Since the optical depth below which convection becomes inefficient is relatively high, the wind of a super Eddington object should be very thick and the mass loss should be enormous if nothing happens above the convective zone (cf Shaviv 2000a).

When the radiative force term becomes as important as the gravity or the gas pressure gradient terms, it can be a source of mechanical types of instabilities. For example, Glatzel & Kiriakidis (1993) have shown that strange modes (which are 'opacity' modified acoustic waves) suffer an instability in which mechanical work by the radiation is pumped into the waves. This instability arises from the interaction of the acoustic waves with the non local nature of the diffusive radiation equations. Another instability that requires special opacity laws is the κ-mechanism. Unlike the s-mode instability, however, it grows on time scales much slower than the dynamical time scale, it does not involve work done by the radiation, and exists only when the system is neither fully adiabatic nor fully in the opposite (NAR) limit. However, it can exist when the system is far from Eddington, because it involves the transfer of heat and not work by the radiation.

A related instability arises when the radiation field transfers energy into acoustic waves due to high radiation pressures, as was shown by Hearn (1972, 1973) and elaborated by Berthomieu et al. (1976), Spiegel (1976), Carlberg (1980) and Asplund (1998). The instability arises when the opacity increases under compression and the change in the opacity is synchronized with the radiative force. This instability is not a real instability (aka 'an absolute instability') in the sense that an infinitesimal perturbation will not grow to become nonlinear, however, as it is a secular drift instability (aka 'a convective instability'), it can take finite amplitude waves (excited by convection for example) and increase their amplitude by many e-folds during one crossing time of the system. (A counter propagating wave will be dissipated as fast, thus adding up to no net amplification of a standing wave unless the conditions for the strange mode instability are satisfied, e.g., Glatzel 1994, Papaloizou et al. 1997). An additional term arises when the Doppler variations of the radiative absorption (arising from wave motion) are taken into account (Shaviv 1999b). A third term is the driving through line absorption. This term which utilizes the Doppler effect as well, is often very important in wind acceleration from hot stars (Owocki & Rybicki 1984, 1985, 1986, 1991, Rybicki, Owocki & Castor 1990, Feldmeier et al. 1997). The modes in systems with this type of simple drift instabilities cannot be used however to compose unstable standing waves in a simple Thomson scattering dominated atmosphere.

The problem of the behavior of luminous atmospheres has also been compared to fluidized beds (Prendergast & Spiegel 1973). Intuitively, the kind of bubbling observed in fluidized beds should also manifest itself in radiatively supported atmospheres. Additional effects arise with the introduction of strong magnetic fields. Specifically, when strong magnetic fields are present such that hydrodynamic motion perpendicular to the magnetic field lines is frozen, a linear analysis shows that 'photon bubbles' arise at a high luminosity, facilitating the transfer of energy (Arons 1992). The original context of accretion onto magnetized

neutron stars was also used to describe accretion disks as well (Gammie 1998). Magnetic fields are neglected in the present work. As we shall see, a plethora of instabilities can already manifest themselves when the magnetic fields are dynamically not important.

In a recent work by Spiegel & Tao (1999), it was shown using a global non-radial linear mode analysis, that atmospheres can have unstable standing modes (as opposed to the secular drift instability, or convective instability) even in the case of pure Thomson scattering (as opposed to strange mode instability). These unstable modes were shown to be horizontally propagating waves that amplify over a time scale much longer than the dynamical times scale of the system. We have tried to recover this finding but could not succeed. We report in the Appendix how such an instability could artificially arise. Moreover, the two other soon be described instabilities are dynamically much more important.

In addition to the intrinsic variability and the generation of winds, perhaps the most interesting consequence of the instabilities is the change induced in the Eddington limit. Shaviv (1998) has shown that in the presence of inhomogeneities, the spatially averaged flux through an atmosphere can increase without increasing the time averaged force on it. This arises because of the tendency of radiation to find the paths of least resistance when escaping through an atmosphere. Under some opacity laws, this will translate into having more radiation in regions where there is less mass and a smaller average force will be exerted. Thus, the luminosity can exceed the Eddington luminosity while the average force on a parcel of mass can remain less than the critical one.

In summary, the classical Eddington limit is restricted to complete homogeneous systems and is frequently not valid in inhomogeneous ones. A fundamental question of how such inhomogeneities are formed and what form they accept is discussed only briefly in Shaviv (1998). Here we expand on this question with emphasis on Thomson scattering dominated atmospheres ('Thomson atmospheres' hereafter).

Evidently, a picture which relates all the different instabilities together and gives sound physical reasoning to their origin is mostly absent even though it is of paramount importance to the understanding of very luminous objects. It is therefore, our goal in this paper to address the following points:

1. We wish to unravel all the absolute instabilities that Thomson atmospheres (which are the simplest of all) exhibit as they approach the Eddington limit. This is done using a rigorous linear analysis (that is similar to but not the same as that of Spiegel & Tao 1999) which assumes nothing save a gray slab atmosphere under the Eddington approximation.

2. We wish to understand the physical basis behind the instabilities found and how systems with these instabilities behave.

Two instabilities will be found and described. Additional instabilities arise when the effects of a varying opacity are taken into account. As the Eddington limit is approached, these instabilities generally become less important as the opacity becomes mostly that of scattering. However, a future 'tabulation' of all the instabilities of luminous atmospheres should include them as well.

A few clear extensions, such as a complete nonlinear analysis using hydrodynamic simulations or a systematic study of the parameter space, will be analyzed in forthcoming publications. These will also include a more elaborate global analysis of the non-trivial opacity law triggered instabilities.

The paper is organized in the following way: The governing radiative hydrodynamic equations are presented in §2. To clarify the physical picture, we discuss in §3 the parameter space of the problem and show where and when the various limits appear. We then proceed in §4 to analyze the problem in a full global analysis assuming only a gray plane parallel atmosphere, calculate the eigenmodes, and find the characteristics of the instabilities found. We continue in §5 to describe the first instability and in §6 the second instability. We end with a discussion and summary of conclusions. An appendix is given to relate the results found here to those found by Spiegel & Tao (1999).

2. THE RADIATIVE HYDRODYNAMIC EQUATIONS AND NOTATION

We begin by laying out the equations used in this work. The equations describing the gas are the continuity, momentum and energy equations:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \tag{1}$$

$$\rho \frac{D\mathbf{u}}{Dt} = -\nabla p - \rho g \widehat{\mathbf{z}} + \rho \frac{\chi}{c} \mathbf{F} \tag{2}$$

$$\frac{Dp}{Dt} - c_a^2 \frac{D\rho}{Dt} = -(\gamma - 1)\rho\kappa c (S - E), \tag{3}$$

with standard notation: p, ρ and $\mathbf{u}$ are the gas (matter) pressure, density and velocity. χ is the total opacity per unit mass, which is the sum of the scattering σ and the absorption κ opacities. Throughout this paper, opacity coefficients without subscripts denote opacity per unit mass while a subscript v denotes opacity per unit volume (with dimensions of $length^{-1}$). S is the equilibrium radiation energy density calculated from the gas temperature $S = 4\sigma_{\rm SB} T^4/c$ where $\sigma_{\rm SB}$ is the Stefan Boltzmann constant. E is the radiation energy density and $\mathbf{F}$ is the radiation energy flux. The radiation energy and flux are those measured in the material rest frame, namely, the 'Lagrangian' quantities. $D/Dt \equiv \partial/\partial t + \mathbf{u} \cdot \nabla$ is the comoving derivative. We distinguish between the adiabatic speed of sound $c_a^2 = \gamma p/\rho$ and the isothermal speed of sound $c_T^2 = p/\rho$. Gravity acts downward along the $\widehat{\mathbf{z}}$ direction.

In view of the above assumptions, we describe the radiation field using the Eddington approximation. Its 'Lagrangian' equations, valid to $\mathcal{O}(v/c)$, are (Mihalas & Weibel Mihalas 1984):

$$\frac{DE}{Dt} - \frac{4E}{3\rho}\frac{D\rho}{Dt} + \nabla \cdot \mathbf{F} = \rho\kappa c (S - E) \tag{4}$$

$$\frac{D\mathbf{F}}{Dt} + \frac{c^2}{3}\nabla E = -\rho(\kappa + \sigma) c \mathbf{F}. \tag{5}$$

The second term in the first equation is the 'compression' term which corrects the stationary frame radiation quantities into the material rest frame. This 'Doppler' term describes radiation drag.

The Eddington approximation is a reasonable assumption since we are interested in regions where $\infty > \tau \gtrsim 1$. Even at the photosphere where the approximation is least accurate, it yields reasonable quantitative results and the correct qualitative results (e.g., Mihalas & Weibel Mihalas 1984, pp. 357, 518).

Once we apply a perturbation, quantities with an index '0' are the unperturbed values while an index '1' indicates the first order perturbation. The first order perturbation to the opacity, $\chi_{v,1}$, can be written as:

$$\chi_{v,1} = \chi_{v,0}\left[\bar{\chi}_\rho \frac{\rho_1}{\rho_0} + \bar{\chi}_T \frac{T_1}{T_0}\right], \tag{6}$$

with the following definitions for the logarithmic derivatives:

$$\bar{\chi}_\rho \equiv \left.\frac{\partial \ln \chi_v}{\partial \ln \rho}\right|_T \quad \text{and} \quad \bar{\chi}_T \equiv \left.\frac{\partial \ln \chi_v}{\partial \ln T}\right|_\rho. \tag{7}$$

The pure Thomson scattering case corresponds to $\bar{\chi}_\rho = 1$ and $\bar{\chi}_T = 0$. However, the approximation we will use in this paper is

$$0 \neq \kappa/\sigma \ll 1, \tag{8}$$

namely, the absorption is very small compared to the scattering, so that the derivatives are to a good approximation those of pure Thomson. However, the absorption is not zero and the gas absorbs a small amount of radiation. This allows at times for the gas to equilibrate its temperature with the radiation, something that for a purely scattering atmosphere cannot take place.

The last equation is the equation of state of a perfect gas relating the gas variables:

$$p = \frac{\rho k_B T}{A m_H}, \tag{9}$$

where A is the mean atomic weight of the gas and k_B the Boltzmann constant.

As the radiation field may not be in equilibrium with the matter, the expression for the opacity can be very complicated and in principle the opacity should be calculated consistently with the radiation field. In the present work we restrict the discussion to a gray opacity.

The Fourier transform of a quantity x is denoted by $\tilde{x}$.

3. DIMENSIONAL ANALYSIS AND CRITICAL PARAMETERS

The driving forces on the system are the terms that appear on the r.h.s. of eq. (2). In equilibrium, the radiative and pressure forces balance the pull of gravity and consequently the r.h.s. vanishes. The relative importance of each term can be related to the ratio between the radiation pressure and the gas pressure:

$$\beta \equiv \frac{p_{\rm rad}}{p} = \frac{E_0/3}{c_T^2 \rho_0}, \tag{10}$$

where $p \equiv p_{\rm gas}$.

A second related parameter which determines the ratio between the radiative and gravitational forces is:

$$\Gamma = \frac{F\chi}{cg}, \tag{11}$$

where F is the radiative flux and χ the opacity per unit mass. Γ is often called the Eddington parameter. If χ is taken to be the Thomson scattering cross-section, then $\Gamma = 1$ corresponds to the Eddington limit. For other opacities (which are generally larger if the gas is ionized) $\Gamma = 1$ corresponds to the Modified Eddington limit.

The parameters β and Γ are simply related to each other in Thomson atmospheres, for which the opacity per unit mass is constant. From the Eddington approximation we have that

$$E_0 = (2 + 3\tau)\frac{F_0}{c}, \tag{12}$$

where τ is the optical depth for extinction. The decrease in the effective gravitational pull $(1-\Gamma)g$ leads to an increase in the density scale height of the atmosphere, namely,

$$l_\rho \equiv \frac{c_T^2}{g(1-\Gamma)}. \tag{13}$$

If the radiative force is constant and isothermality is assumed, then the puffed up scale height is constant with height. In the case of an atmosphere with an adiabatic gradient, c_T should be replaced with c_a, in which case the scale height is not constant with height any more. To demonstrate the relation between β and Γ we assume for simplicity an isothermal photosphere. In this approximation,

$$\begin{aligned} \rho_0(z) &= \rho_0(z=0)\exp(-z/l_\rho); \\ p_0(z) &= c_T^2 \rho_0(z=0)\exp(-z/l_\rho). \end{aligned} \tag{14}$$

The optical depth from $z = \infty$ to a given point at a depth of z is

$$\tau = \int_z^\infty \rho\chi dz = \chi l_\rho \rho_0(z=0)\exp(-z/l_\rho) = \chi l_\rho p_0(z)/c_T^2. \tag{15}$$

For further reference we also have $\tau = \chi_v l_\rho$. We thus find from eqs. (10) through (15) that

$$\begin{aligned} \beta \equiv \frac{p_{\rm rad}}{p} &= (2+3\tau)\frac{l_\rho F_0 \chi}{4c\tau c_T^2} = \frac{(2+3\tau)}{3\tau}\frac{\Gamma}{1-\Gamma} \\ &\approx \frac{\Gamma}{(1-\Gamma)} \text{ for } \tau \gg 1. \end{aligned} \tag{16}$$

Conversely, for an optically thick atmosphere, we have that

$$\Gamma \approx \frac{\beta}{\beta+1}. \tag{17}$$

Atmospheres that have a small radiation to gas pressure ratio will therefore have a luminosity that is considerably smaller than the Eddington luminosity (given by $\Gamma = 1$). When the gas and radiation pressures equate, we have that $\Gamma \approx 1/2$. When the radiation pressure is much larger than the gas pressure, Γ approaches unity (such that $1-\Gamma \ll 1$). In general, when the opacity is a function of height or when the non-isothermalness of the atmosphere is considered, we will get a more complicated relation between Γ and β, however, we will still find that $\beta \ll 1$ corresponds to $\Gamma \ll 1$, that $\beta \approx 1$ corresponds to $\Gamma \sim 1/2$ and that $\beta \gg 1$ corresponds to $(1-\Gamma) \ll 1$.

Additionally important parameters arise when comparing the different time scales found in the problem. The first time scale is the dynamic time scale defined here as the isothermal sound crossing time of a density scale height:

$$\tau_{\rm dyn} \equiv \frac{l_\rho}{c_T} = \frac{c_T}{(1-\Gamma)g}. \tag{18}$$

From analysis of equations (4) and (5), a second time scale appears. To see this, we combine the two linearized versions of the equations assuming the heat absorption is negligible (i.e., in the adiabatic or isothermal limits) and neglecting the 'Doppler' term, to get:

$$\frac{1}{c^2}\frac{\partial^2 E}{\partial t^2}+\frac{\chi_v}{c}\frac{\partial E}{\partial t}-\frac{1}{3}\nabla\cdot\nabla E=0, \tag{19}$$

where $\chi_v=\rho\sigma$. In the more general case the coefficient of $\partial E/\partial t$ is $\rho(\sigma+\kappa)/c$. If the typical length scale is given by k^{-1}, a diffusion time scale can be defined as

$$\tau_{\rm diff}(k)\equiv\frac{3\chi_v}{ck^2}, \tag{20}$$

when comparing the second and third terms. Specifically, for a wavelength of the order of the scale height of the atmospheres, we can define:

$$\tau_{\rm diff}\equiv\tau_{\rm diff}(k=2\pi l_\rho^{-1})\equiv\frac{3}{(2\pi)^2}\frac{\chi_v l_\rho^2}{c}. \tag{21}$$

A dimensionless parameter is its ratio to the dynamical time scale, given by

$$\begin{aligned} r_{\rm diff}&\equiv\frac{\tau_{\rm diff}}{\tau_{\rm dyn}}=\frac{3}{(2\pi)^2}\frac{\chi_v l_\rho^2}{c}\frac{c_T}{l_\rho} \\ &=\frac{3}{(2\pi)^2}\left(\frac{c_T}{c}\right)(\chi_v l_\rho)\equiv\frac{3}{(2\pi)^2}\left(\frac{c_T}{c}\right)\tau_0, \end{aligned} \tag{22}$$

where $\tau_0\equiv\chi_0 l_\rho$ is the typical optical depth of a density scale height. For $r_{\rm diff}\ll 1$, the radiation has ample time to converge to the diffusive solution given by the instantaneous mass configuration. In the opposite limit, the radiation does not have enough time to diffuse and any perturbation to it is synchronized with the gas.

Note that for the optically thin part of the atmosphere, one should consider the nonlocal nature of the diffusive solution. In such a case, the diffusive time scale is simply given by the time it takes the radiation to traverse a scale height, i.e., $\tau_{\rm diff}\approx l_\rho/c$, yielding $r_{\rm diff}\approx c_T/c$.

A third time scale emerges from the energy equation (eq. 3). We look at the set of eqs. (1)-(5) in the linearized case, when zero order gradients are neglected and when the terms arising from the finite propagation speed of light can be neglected as well. After applying a Fourier transformation, eqs. (4) & (5) become

$$\tilde{S}_1-\tilde{E}_1=\tilde{S}_1\left(\frac{(k^2/3\chi_v\kappa_v)}{1+(k^2/3\chi_v\kappa_v)}\right), \tag{23}$$

where χ_v is the total extinction per unit volume and $\tilde{S}_1$ and $\tilde{E}_1$ are the Fourier transforms of S_1 and E_1 respectively. We find that for $(k^2/3\chi_v\kappa_v)\gg 1$, $\tilde{S}_1-\tilde{E}_1\approx\tilde{S}_1$ while for $(k^2/3\chi_v\kappa_v)\ll 1$, $\tilde{S}_1-\tilde{E}_1\approx\tilde{S}_1(k^2/3\chi_v\kappa_v)$. If we now plug this result into the linearized and Fourier transformed energy equation (eq. 3) and use the first two hydrodynamic equations (eqs. 1 & 2) to simplify it, we obtain after some algebra (in a similar manner to §101 of Mihalas & Weibel Mihalas 1984) that

$$\left\{\left[\omega^2-c_a k^2\right]\omega+\tau_{\rm cool}^{-1}(k)\left[\omega^2-c_T k^2\right]\right\}\tilde{p}=0, \tag{24}$$

where

$$\tau_{\rm cool}(k)\equiv\frac{1}{12\beta(\gamma-1)\kappa_v c}\left[1+3\frac{\chi_v\kappa_v}{k^2}\right]. \tag{25}$$

Clearly, when $\tau_{\rm cool}^{-1}\gg\omega$, the wave equation obtained will be the isothermal wave equation while for $\tau_{\rm cool}^{-1}\ll\omega$, the adiabatic wave equation is obtained. $\tau_{cool}(k)$ is therefore, the typical time scale which takes a wave with a wavevector k to cool radiatively.

We can define a dimensionless ratio between the cooling time scale for a wave with a wavelength of the order of the scale height and dynamical time scale that is the time it takes the same wave to cross a scale height as:

$$\begin{aligned} r_{\rm cool}&\equiv\tau_{\rm cool}(k=2\pi l_\rho^{-1})/\tau_{\rm dyn} \\ &=\frac{c_T}{12\beta(\gamma-1)c\kappa_v l_\rho}\left(1+\frac{3\chi_v\kappa_v l_\rho^2}{(2\pi)^2}\right) \qquad (26) \\ &\approx\begin{cases}\frac{1}{12\beta(\gamma-1)}\frac{c_T}{c}\frac{1}{\tau_0}\frac{\chi_v}{\kappa_v} & \text{for}\quad \chi_v\kappa_v l_\rho^2\ll 1\\ \frac{1}{16\pi^2\beta(\gamma-1)}\frac{c_T}{c}\tau_0 & \text{for}\quad \chi_v\kappa_v l_\rho^2\gg 1.\end{cases} \qquad (27) \end{aligned}$$

For $r_{cool}\gg 1$, the cooling time scale is too long for the perturbations on scales of the scale height to have time to equilibrate to the radiation temperature. The gas is then close to the adiabatic limit. In the opposite limit of $r_{cool}\ll 1$, the gas has time to reach thermal equilibrium with the radiation and if the radiation has no perturbation, it is close to the isothermal limit. Since the equilibrium of the radiation can in principle be non-isothermal (if there is a net flux in the system), this limit is more generally called the NAR limit (non-adiabatic reversible), since in this opposite limit to adiabatic perturbations, the perturbations are again reversible.

Fig. 1 depicts the typical value of the dimensionless parameters in a typical Thomson atmosphere as a function of the optical depth from the top of it. Four distinct domains are seen in the $\tau_{\rm cool}/\tau_{\rm dyn}$ vs. $\tau_{\rm diff}/\tau_{\rm dyn}$ plane. The various approximations are depicted in this plane. The figure illustrates the run of the physical conditions in a typical star as one moves from the inside outwards, namely from high to low optical depths. The effect of the ratio of extinction to absorption on the path in the parameter plane is also shown. The discussion in this paper concentrates on the l.h.s. of the parameter plane, that it, on the region in which the radiation does not evolve synchronously with the instantaneous gas configuration, i.e., when the radiation and gas cannot be considered as one fluid.

4. FULL GLOBAL LINEAR ANALYSIS

We begin with a complete global linear analysis. By changing the free parameters of the problem (Γ, boundary conditions, etc.) we expect to find the various instabilities that exist in this type of systems. Afterwards, we will analyze each one of them separately.

4.1. *Equations and Problem set-up*

The governing set of equations is given in §2. After linearization and Fourier transforming the *horizontal* direction and *time* we get the following equations for the transformed quantities:

$$i\omega\frac{\tilde{\rho}}{\rho_0}=ik\tilde{u}+\frac{\partial}{\partial z}\tilde{w}+\tilde{w}\frac{1}{\rho_0}\frac{\partial}{\partial z}\rho_0 \tag{28}$$

$$\begin{aligned} -i\omega\rho_0\tilde{w}=&-\frac{\partial}{\partial z}\tilde{p}-g_*\tilde{\rho}+\rho_0\frac{(\kappa+\sigma)}{c}\tilde{F}_z \qquad (29)\\ &+\rho_0\left(\frac{(\kappa+\sigma)}{c}\right)F_0\left[\chi_\rho\frac{\tilde{\rho}}{\rho_0}+\chi_p\frac{\tilde{p}}{p_0}\right] \end{aligned}$$

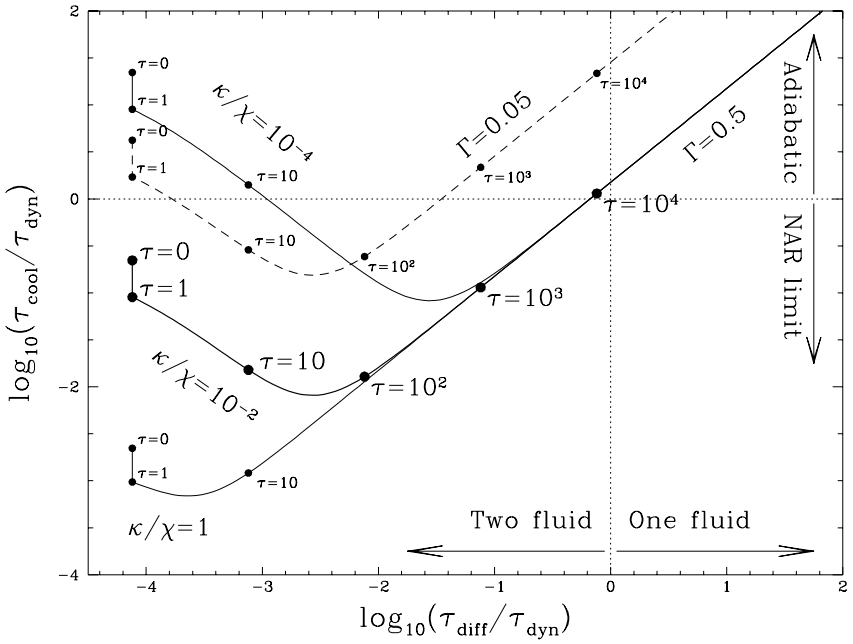

FIG. 1.— The connection between the ratio of the cooling time scale to the dynamical time scale and the ratio of the diffusion time scale to the dynamical time scale for a few sample Thomson atmospheres, as a function of optical depth. The case depicted by a thick solid line corresponds to $\gamma = 5/3$, $c_T/c = 10^{-3}$, $\kappa/\chi = 10^{-2}$ and $\Gamma = 0.5$ ($\beta = 1$ in the optically thick limit). The two lighter solid lines correspond to changing the absorptive opacity from $\kappa/\chi = 10^{-2}$ to 1 or to 10^{-4}. The dashed line corresponds to changing Γ from 0.5 to 0.05. The four regions correspond to whether the radiation is in the diffusion limit or not (i.e., whether it behaves as two fluids or one coupled fluid), and whether the gas is in the adiabatic or NAR limits. Convection is effective in the top half. The instabilities described here operate in the left half.

$$-i\omega\rho_0\tilde{u} = -ik\tilde{p} + \rho_0\frac{(\kappa+\sigma)}{c}\tilde{F}_x \tag{30}$$

$$-i\omega\tilde{E} + \frac{\partial}{\partial z}\tilde{F}_z = \rho_0\kappa c(\tilde{S}-\tilde{E}) - \frac{4E_0}{3\rho_0}i\omega\tilde{\rho} - ik\tilde{F}_x \tag{31}$$

$$-i\omega\tilde{F}_x + ik\frac{c^2}{3}\tilde{E} = -\rho_0(\kappa+\sigma)c\tilde{F}_x \tag{32}$$

$$-i\omega\tilde{F}_z + \frac{c^2}{3}\frac{\partial}{\partial z}\tilde{E} = -\rho_0(\kappa+\sigma)c\tilde{F}_z - \tilde{\rho}(\kappa+\sigma)cF_0 \tag{33}$$

$$-\rho_0(\kappa+\sigma)cF_0\left[\chi_\rho\frac{\tilde{\rho}}{\rho_0} + \chi_p\frac{\tilde{p}}{p_0}\right]$$

$$-i\omega\tilde{p} - \rho_0 g_*\tilde{w} = -c_s^2\rho_0(ik\tilde{u} + \frac{\partial}{\partial z}\tilde{w}) \tag{34}$$

$$-(\gamma-1)\rho_0\kappa c(\tilde{S}-\tilde{E})$$

$$\tilde{S} = 4E_0(\frac{\tilde{p}}{p_0} - \frac{\tilde{\rho}}{\rho_0}), \tag{35}$$

where w is the vertical component of the velocity and u the lateral one. These equations are the equations that appear in (Spiegel & Tao 1999) with two differences. First, we write the radiation in the co-moving system with the material. Second, we retain the option of a variable opacity for generality and usage in the next work. We omit for brevity the '1' index from the Fourier transformations of our eight variables $\tilde{\rho}_1$, $\tilde{p}_1$, $\tilde{u}_1$, $\tilde{w}_1$, $\tilde{E}_1$, $\tilde{S}_1$, $\tilde{F}_{x,1}$ and $\tilde{F}_{z,1}$. These variables satisfy four complex algebraic equations and four complex ODE's with the height z as the independent variable.

The Boundary Conditions: The set of ODE's requires boundary conditions. Since the dependent variables $\tilde{p}, \tilde{w}, \tilde{E}, \tilde{F}_z$ are complex variables, four complex boundary conditions are required.

Several gas boundary conditions can be imposed. At either the bottom or the top we choose $w = 0$ or $p = 0$, which result with reflection of the waves at the boundary. We find that these conditions do not change the qualitative behavior of the waves. We should note however, that if we wish to compare our results to the behavior of real atmospheres, the results will have any meaning only if the waves in the real atmosphere considered can be trapped (and not for example propagate outward where they can be dissipated). We shall elaborate this point more in the discussion section.

As for the radiation, there are more possibilities. The following possibilities exist for the **top**:

Fixed radiation temperature: $\tilde{E}(z_t) = 0$ or

Eddington $\tau = 0$ condition: $\tilde{E}(z_t) = 2\tilde{F}(z_t)/c$

While the following possibilities exist for the **bottom**:

Fixed radiation temperature: $\tilde{E}(z_b) = 0$ or

Fixed incident flux: $\tilde{F}(z_b) = 0$

When the top of an atmosphere is free to emit radiation to infinity, the Eddington condition for the top is the more appropriate condition. The fixed radiation temperature will be the physically appropriate one if for example the layer of interest is bounded at the top by a layer that equilibrates quickly its temperature by convection. As to the bottom conditions, they too can depend on the system at hand. For example, if the atmosphere is above a layer which conducts heat well, its temperature will be fixed. In principle, this boundary condition can be a mix of the two.

4.2. *Form of Solution*

In order to find the eigenmodes of the acoustic oscillations, we first need to construct a profile of the unperturbed atmosphere. We avoid at this stage the usage of real stellar atmospheres since we are more interested here in understanding the physical origin and behavior of the unstable modes in the simple Thomson scattering dominated case and not to quantify the unstable behavior of real stars that have complex atmospheric opacities. More realistic opacities as well as non LTE effects will be included in future work. As stated before, we also neglect for simplicity the non-slab geometry of most relevant atmospheres.

The equations describing the unperturbed steady state atmosphere are the radiation diffusion equations:

$$F_0 = const \tag{36}$$

$$\frac{dE_0}{dz} = -\frac{3\chi_{v,0}F_0}{c}, \tag{37}$$

and the hydrostatic equation

$$\frac{dp_0}{dz} = c_T^2\frac{d\rho_0}{dz} = -\rho_0 g\hat{\mathbf{z}} + \frac{\chi_{v,0}F_0}{c}. \tag{38}$$

Since the gas and radiation are in thermal equilibrium in the unperturbed solution, we also have $T^4 = cE/4\sigma_{\rm SB}$ which allows the calculation of c_T.

Inasmuch as the Eddington relation should be satisfied in particular in the optically thin limit, we require that at the top of the atmosphere $E(\tau \to 0) = 2F/c$.

The main parameter of interest in this work is $\Gamma = \chi_{v,0}F/cg$. Hence, we solve for an unperturbed atmosphere with a given Γ extending over a given optical depth range (say, between $\tau = 0.1 - 100$). The solution is carried out numerically and then tabulated so that a simple spline interpolation can be applied. The values of p, ρ or E can be quickly and to any given prescribed accuracy be found for any height in the atmosphere.

The next step is to find the eigenmodes of waves with a given horizontal wavevector k_x in a given atmosphere. Our set of ODEs has four first order complex ODE's and complex eigenvalues. Two of the boundary conditions are given at the bottom of the atmosphere while the two others are given at the top, making it a complicated boundary value problem. To find an eigenmode we use the shooting method. The two 'missing' bottom boundary conditions are iterated for until the imposed top boundary conditions are satisfied. The procedure of finding the solution is carried out in the following way:

1. Since two boundary conditions are given at the bottom, we guess the values of the two other variables that don't have a specified value on the boundary along with a guess for the unknown eigenfrequency ω. For example, if $\tilde{p} = 0$ and $\tilde{E} = 0$ are the bottom boundary conditions, we start by guessing the values of $\tilde{F}_z$, $\tilde{w}$ and ω.

2. Once all the dependent variables are given or guessed on the lower boundary, we can integrate upwards. The set of equations can be written as four ODE's and four algebraic equations that can be evaluated at each point z in the following order:

   - Evaluate $\kappa, \sigma, \chi = \kappa + \sigma, E_0, s_\rho = -{\rho_0}^{-1}(\partial\rho_0/\partial z)$ and $g_* = g - (\chi/c)F$.
   - Calculate numerically the variables with algebraic equations:

$$\tilde{F}_x = \frac{kc^2\tilde{E}}{3(u_c\omega + i\rho_0\chi c)} \tag{39}$$

$$\tilde{u} = \frac{k\tilde{p}}{\omega\rho_0} + \frac{i}{\omega}\frac{\chi}{c}\tilde{F}_x \tag{40}$$

$$\tilde{S} = \left[i\omega\tilde{p} - c_s^2\rho_0(i\omega\frac{\tilde{p}}{p_0} + s_\rho\tilde{w}) + \tilde{E}(\gamma-1)\rho_0 c\kappa_0 u_\kappa^{-1} + \rho_0 g_*\tilde{w}\right] / \left[(\gamma-1)\rho_0 c\kappa_0/u_\kappa - i\omega {c_s}^2/4E_0\right] \tag{41}$$

$$\tilde{\rho} = \rho_0\left(\frac{\tilde{p}}{p_0} - \frac{\tilde{S}}{4E_0}\right). \tag{42}$$

   - Integrate upwards using the following ODE's:

$$\frac{\partial}{\partial z}\tilde{w} = i\omega\frac{\tilde{\rho}}{\rho_0} - ik\tilde{u} + \tilde{w}s_\rho \tag{43}$$

$$\frac{\partial}{\partial z}\tilde{p} = i\omega\rho_0\tilde{w} - g_*\tilde{\rho} + \rho_0\frac{\chi}{c}\tilde{F}_z + \rho_0 F_0\frac{\chi}{c}\left[\chi_\rho\frac{\tilde{\rho}}{\rho_0} + \chi_p\frac{\tilde{p}}{p_0}\right] \tag{44}$$

$$\frac{\partial}{\partial z}\tilde{F}_z = u_\kappa\rho_0\kappa c(\tilde{S} - \tilde{E}) + u_c i\omega\tilde{E} - ik\tilde{F}_x - i\omega\frac{4E_0}{3\rho_0}u_{\rm d} \tag{45}$$

$$\frac{\partial}{\partial z}\tilde{E} = -3\rho_0\frac{\chi}{c}\tilde{F}_z + \frac{u_c 3i\omega}{c^2}\tilde{F}_z - 3\tilde{\rho}\frac{\chi}{c}F_0 - 3\rho_0\frac{\chi}{c}F_0\left[\chi_\rho\frac{\tilde{\rho}}{\rho_0} + \chi_p\frac{\tilde{p}}{p_0}\right]. \tag{46}$$

   We marked several terms so as to be able to trace their effect. The terms multiplied by u_c are terms that originate from the wavy behavior of the radiation field, while terms multiplied by u_κ are terms that originate from the finite cooling time of the system. The terms with $u_{\rm d}$ are the 'Doppler' terms. In the normal case $u_c = u_\kappa = u_{\rm d} = 1$. In the limit $u_c \to 0$ we get the instantaneous limit for the radiation. While in the limit $u_\kappa \to 0$ we recover the negligibly small gas heat capacity limit (the NAR limit). The limit $u_{\rm d} \to 0$ shuts off the effects of the Doppler correction in the radiation field to the motion of the material. Since one of the goals is to check whether these terms are important for instability, they are specifically marked and can be artificially changed, though they are equal to unity in real cases.

   The numerical integration is accomplished using the fourth order Runge-Kutta integration scheme.

3. At the top, the solution is compared with the two top boundary conditions. The initial guess for the two additional bottom conditions and the eigenfrequency ω are then corrected and the equations are integrated upwards once again. The 'shooting' upwards it iterated until the solution converges to the required top boundary conditions. The solution obtained is an eigenmode of the system and its eigenvalue is the converged frequency ω.

Since we are interested in finding more than one particular eigenmode, we use the above procedure to search for more modes. This search is performed by systematically guessing different initial guesses for the eigenvalue ω. In this way, the entire frequency domain is covered up to a given frequency.

In principle, other types of modes such as radiation diffusion modes at very high frequency can be searched for and found, however, they are beyond the interest of this work as they are generally very stable (with large damping rates, e.g. Mihalas & Weibel Mihalas 1984). Similarly, we are not interested in convection or g-modes. For any reasonable absorptive opacity larger than $\sim (c_a/c)(p_{\rm rad}/p_{\rm gas})\sigma$, the top part of the atmosphere will cool too quickly for convection to be efficient. This will appear in the form of modes that have are real (g-modes) or pure imaginary (convective modes) with absolute oscillation or growth rates that are much smaller than the dynamical time scale. Only if we force the atmosphere to be adiabatic by artificially having a very small absorptive opacity, will the g-modes or convective modes obtain a 'dynamic' time scale.

Once the eigenmodes of waves with a given k_x and Γ are found, we can systematically search for the eigenmodes of other k_x, or of atmospheres with a different Γ. Since we are interested in the locus of the eigenvalues as a function of either k_x or Γ, we can evaluate the locus by changing k_x or Γ by small increments and use the previous value of the two independent variables and eigenvalue as the new initial guess. We assume that the eigenvalues are continuous functions of k_x and Γ. This allows a relatively fast construction of the locus of eigenvalues.

4.3. *Results*

We divide the description of the results into three parts. We first begin by analyzing the eigenvalues ω of an atmosphere with a given Γ and horizontal wavenumber k_x and summarize the two instabilities found. We proceed to construct the eigenvalue spectrum of an atmosphere as a function of the horizontal wavevector k_x and then perform a similar analysis to find the eigenvalues as a function of the Eddington parameter Γ. In subsequent sections, we will then study the two absolute instabilities.

4.3.1. *An atmosphere with a given Γ and k_x and the instabilities found*

As described in §4.2, for a given set of boundary conditions imposed, the eigenvalues corresponding to a given Γ, k_x pair are found.

Fig. 2 depicts the lowest acoustic eigenvalues obtained for three atmospheres with $\Gamma = 0.9$ and $k_x = 1/l_\rho$. One atmosphere has a fixed temperature imposed at the bottom boundary condition while the other two have a fixed flux imposed at the bottom. In all three cases the Eddington approximation condition is imposed at the top (namely, $E_t = 2F_t/c$), at $\tau_t = 0.3$. In the fixed temperature case, the bottom is fixed at $\tau_b = 10$. In the two fixed flux atmospheres, the bottom is located at $\tau_b = 10$ and 30.

We assume that there is a small absorptive opacity $\kappa = \chi/100$. This absorptive opacity, however small, is still 100 times larger than the value needed to force the perturbations with a period of the dynamical time scale to be in the NAR limit and not the adiabatic one. This is a better approximation for stellar atmospheres that most often have a small but significant absorptive opacity in this respect. One the other hand, the absorptive opacity is too small to change the total opacity which is dominated by electron scattering. The gravitational acceleration in this particular case is $10^8 cm\ s^{-2}$, an acceleration which corresponds to a White Dwarf envelope. The physics obtained is to a large extent not a function of this parameter. Choosing a different gravitational constant will just imply that the typical time scale will be different and the effective temperature needed to sustain a given Γ is different as well. In this particular case, the effective temperature obtained is 5.82×10^5 °K. At this temperature, the ratio c_T/c at the center of the layer is roughly 1.0×10^{-4}. (In reality these conditions are similar to those found in novae.)

We can now see the two main instabilities. When the lower boundary condition has a fixed temperature, the lowest vertical mode can be unstable to the first type of instability – 'Type I'. When unstable, this mode's eigenfrequency becomes purely imaginary. It grows on dynamical time scales while its counterpart decays with the same rate. The growth rate found in this case is $7.7 s^{-1}$. This is the dynamical time scale of the system.

Under the two different lower boundary conditions, another type of instability can operate – 'Type II'. When a mode succumbs to this instability, is has a real part with a somewhat smaller but significant positive imaginary part. Namely, it is mostly traveling but it amplifies (and its counterpart decays) on a dynamical time scale. In the case depicted here $\omega = (8 + 0.45i)\ s^{-1}$ with only small changes between the three cases. Although the growth rate of this mode is not as fast as fast as that of the Type I instability, the e-folding growth time of this mode is nevertheless only three oscillation periods. Moreover, unlike the Type I instability, this one is relatively insensitive to the boundary conditions.

Fig. 3 shows the structure of both unstable and stable modes. The main features seen are that the Type I mode has gas and radiation energy densities that are inversely correlated with each other. In the Type II mode, there is clearly a phase lag between the density and the radiation. This lag flips sign in the conjugate pair of the Type II unstable mode.

We also checked whether the various artificial factors that we introduced change the occurrence of the modes. u_c changes the ratio between the speed of light and speed of sound. We found that both instabilities exist in the two fluid limit (in which the radiation speed is infinite) but disappear when in the one fluid limit, when the radiation cannot diffuse. u_κ changes the importance of the absorption opacity which determines the ratio between the cooling time scale and the dynamical time scale. We found that both instabilities exist in both limits, the adiabatic limit in which the gas has no time to equilibrate with the radiation temperature and in the NAR limit, in which the gas temperature is set by the radiation temperature.

We also searched for a third instability like the one described by Spiegel & Tao (1999) but could not find it. We defer further discussion on it to the appendix where we will try to understand how such an instability could arise

artificially. We also explain why the two dynamic instabilities found here were not found by Spiegel and Tao even though they did solve the same equations.

4.3.2. *Eigenmodes as a function of k_x*

Taking the two examples depicted in Fig. 2 with $\tau_b = 10$, we now look at the locus of eigenmodes obtained as a function of k_x. This is depicted in Fig. 4 for the fixed temperature case, which has both Type I and Type II unstable modes, and in Fig. 5 for the fixed flux case, which has only Type II unstable modes.

The Type I modes are found to be unstable for all k_x smaller than a critical wavenumber. This wavenumber decreases as Γ is reduced. That is to say, once the critical $\Gamma_{\rm crit} = 1/2$ for instability is surpassed, it is the longest wavelengths which first become unstable. However, since $\Im(\omega)$ is found to be proportional to k_x at long wavelengths, the most unstable mode has a finite horizontal wavelength. This is seen in the additional lines in Fig. 4 which are labeled with varying Γ's. The additional lines describe the growth rate of the Type I unstable mode for different Γ's with the most unstable horizontal wavelength marked with a small open circle.

The Type II modes are seen to occur only in finite wavelengths. In Fig. 5, we can actually see the second eigenmode become Type II unstable in one region of k_x by merging with the third eigenmode and in another region, by merging with the first eigenmode. As the Γ is reduced towards the critical Γ needed for instability, we see that the most unstable mode does not vary by much. The instability is triggered at a finite horizontal wavelength. It is found to be triggered at $\Gamma_{\rm crit} \approx 0.86$ and a wavenumber of $k_{x,\rm crit} \approx 0.88 \left\langle l_\rho^{-1} \right\rangle$. These numbers are found to vary by a few percent if either the absorption opacity or the depth of the atmosphere are changed.

4.3.3. *Eigenmodes as a function of Γ*

Last, we study the behavior of an atmosphere as a function of the Eddington parameter Γ. We take the example depicted in Fig. 4 and instead of changing k_x, we keep it fixed at $k_x = \left\langle l_\rho^{-1} \right\rangle$ and change the Eddington parameter from $\Gamma = 0.1$ to $\Gamma = 0.98$.

It is clear from Fig. 6 that Thomson scattering dominated atmospheres become unstable at large enough luminosities. However, the critical luminosity is well below the Eddington limit. They range from 0.5 in the fixed temperature case to about 0.86 when a fixed flux is imposed at the bottom.

The main feature apparent in the figure is that as the radiative flux increases, more and more modes become unstable to the Type II instability. That is to say, it isn't only the lowest or second lowest vertical eigenmodes which are unstable. The dynamic state that this atmosphere will reach obviously depends on the nonlinear behavior of the modes and their mutual interaction. It is clear however that it will not stay homogeneous close to the Eddington limit.

5. THE TYPE I INSTABILITY

The two main clues pointing to the origin of the Type I instability were seen in the previous sections. First, is was found that the instability arises only when the bottom boundary condition has a fixed temperature, as opposed to a fixed flux. Second, when it does arise, there is clearly an anti-correlation between the radiation energy density and the material density. We now proceed to show how this instability can arise with a toy model.

We begin by showing how the boundary conditions affect the correlation between gas and radiation. We will then proceed to show how an anti-correlation results with the Type I instability.

5.1. *Non local effects arising from a global analysis*

Since the boundary conditions were found to be crucial, we evidently cannot study the instability with a local analysis. We begin with an atmosphere in the NAR limit. Thus, at any given instant, the radiation field is governed by the equilibrium diffusion solution:

$$\nabla \cdot \mathbf{F} = 0 \quad \text{and} \quad \mathbf{F} = \frac{c \nabla E}{3 \chi \rho}, \tag{47}$$

and the boundary conditions imposed at the bottom and top of the atmosphere.

To demonstrate how a positive or negative correlation can arise between E_1 and ρ_1, we explore a specific example. Suppose we have an unperturbed atmosphere in which $\chi_0 = const$ that is perturbed by $\delta\chi_v = \delta\chi_v(x)$, i.e., that the perturbations are only a function of the horizontal direction x. We also assume for simplicity that the vertical variations in the atmosphere can be neglected at this point. This implies that the flux $\mathbf{F}$ and its perturbation do not depend on the vertical coordinate.

The optical depth between the top and any given point in the atmosphere is

$$\tau = \int_z^{z_t} \chi_v dz = \tau_0 \left(1 + \delta\chi_v(x)/\chi_{v,0}\right), \tag{48}$$

where $\tau_0(z)$ is the unperturbed optical depth from infinity down to the mass element at z.

Next, we need to impose boundary conditions. If we impose $\tau = 0$ at the top of the atmosphere, then from the Eddington approximation we have that at the top $E_t = 2F_t/c$. At the bottom we can impose a fixed temperature and therefore a fixed E_b, or, we can impose a fixed flux F_b. An example for the implementation of a fixed temperature (or fixed E_b) at the bottom is the case of a very good energy conducting layer below the layer under consideration.

The fixed flux boundary condition at the bottom corresponds to the case when the layer below the one under consideration has a very large heat transfer resistivity so that the gradient in the layer sets a fixed flux that cannot be changed by the atmosphere exhibiting the instabilities.

When the temperature and energy density E_b are chosen as fixed on the lower boundary, the flux F and the energy density at any height are found to be:

$$\begin{aligned} F = \frac{cE_b}{2+3\tau_b} &= const \\ E = \left(\frac{2+3\tau}{2+3\tau_b}\right) E_b &= \left(\frac{2+3\tau\left(1+\delta\chi_v(x)/\chi_{v,0}\right)}{2+3\tau_b}\right) E_b \\ &\approx E_0(\tau_0)\left(1 - \frac{6(\tau_b-\tau_0)}{(2+3\tau_b)(2+3\tau_0)}\frac{\delta\chi_v(x)}{\chi_{v,0}}\right), \end{aligned} \tag{49}$$

where $E_0(\tau_0)$ is the radiation energy density at the point where the unperturbed optical depth was τ_0. This result

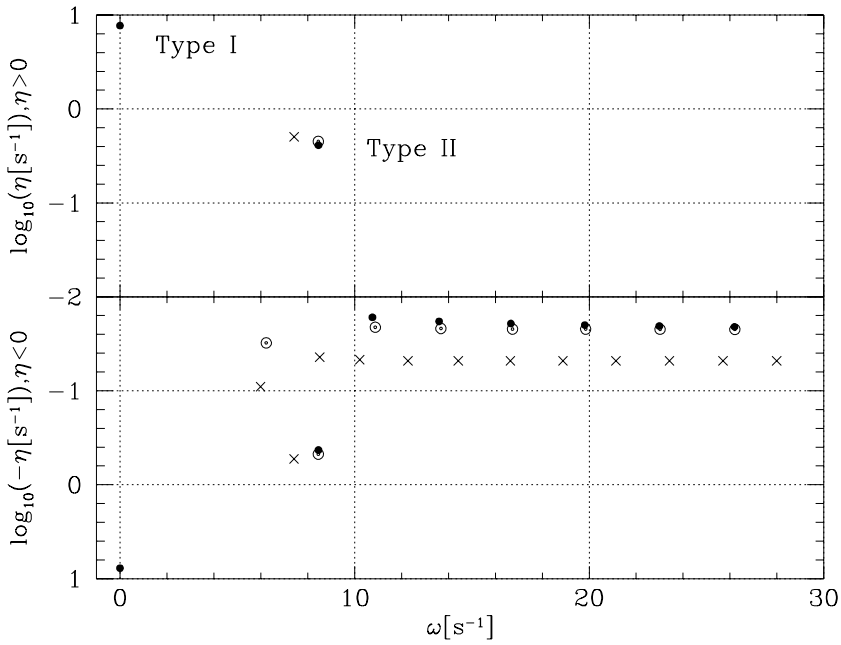

FIG. 2.— The eigen frequencies of the lowest modes in three atmospheres. The two panels refer to the two possibilities of the sign of η – the growth or damping rate of the modes. The filled circles are the modes of an atmosphere with $\Gamma = 0.9$, $k_x = 1/l_\rho$, $g = 10^8 cm\ s^{-2}$, $\kappa/\chi = 1/100$ and a fixed temperature imposed at a bottom placed at $\tau_b = 10$ while the top is at $\tau = 0.3$. The empty circles are the modes of the same atmosphere with a fixed flux imposed at the bottom. The crosses are the eigenmodes obtained when the bottom boundary of the second atmosphere is moved to $\tau = 30$. The Type I instability is seen in only the first atmosphere with a fixed temperature at the bottom. The Type II instability is seen in all cases.

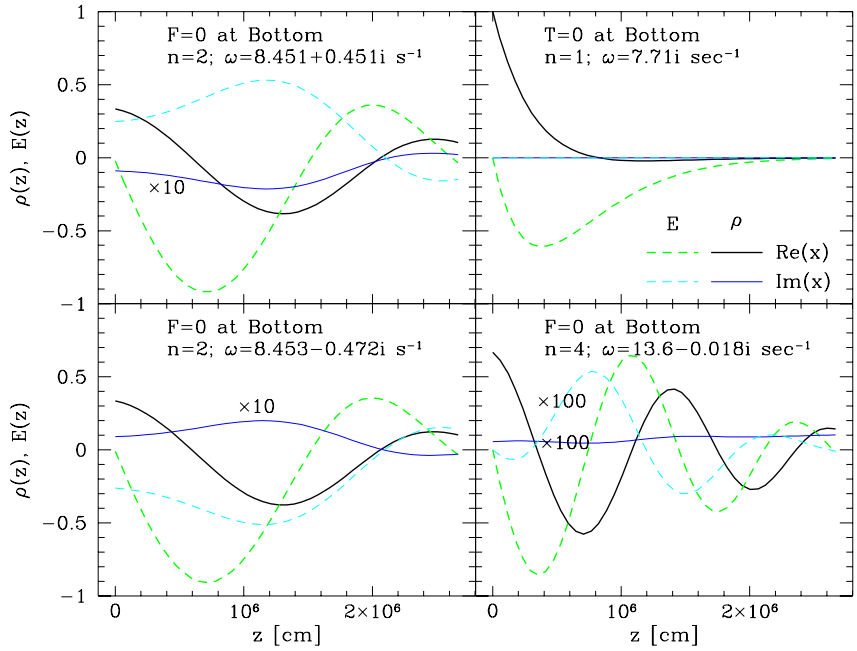

FIG. 3.— The profiles of unstable and stable modes. The top left panel describes the Type II unstable mode of the second atmosphere of the previous figure (fixed flux at the bottom). The bottom left panel depicts the profile of the conjugate mode which is highly damped. The top right panel shows the profile of the Type I unstable mode of the first atmosphere described in the previous figure (fixed T at the bottom). The bottom right panel shows the profile of the fourth mode in the same atmosphere. It is slightly radiation damped. The variables plotted are the energy density $\tilde{E}$ (dashed lines) and the matter density $\tilde{\rho}$ (solid lines) multiplied by βc_T^2, so as to be on the same par as $\tilde{E}$. Since the problem is linear, the absolute normalization of the modes is meaningless. The real part is given by heavy line while the imaginary part by a thin one.

implies that under a positive perturbation to the opacity, the change $\delta E(\tau_0)$ in the radiation energy density will be negative:

$$\frac{\delta E(\tau_0)}{E} \approx -\frac{6(\tau_b - \tau_0)}{(2+3\tau_b)(2+3\tau_0)} \frac{\delta \chi_\nu(x)}{\chi_{\nu,0}}; \qquad (50)$$

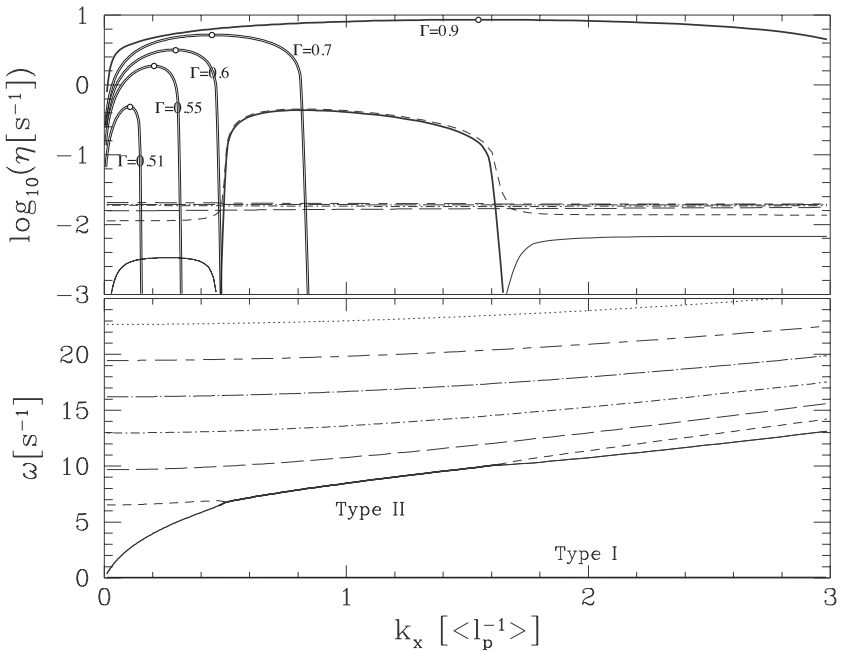

FIG. 4.— The eigenvalues of the first atmosphere depicted in Fig. 2 that has a fixed temperature imposed at its bottom, as a function of the horizontal wavenumber k_x in units of the pressure scale height. The bottom panel is the real part of the eigenmode while the top panel is the $\log_{10}$ of the absolute of the imaginary part. Unstable modes are denoted by thick lines. This atmosphere exhibits both Type I and Type II unstable modes. Below the critical horizontal wavenumber, the Type I unstable mode has a purely imaginary eigenvalue. Above the critical horizontal wavenumber, the absolute value of the real part is much larger than the absolute value of the imaginary part namely, there is a small damping term arising from radiation diffusion. The additional double lines represent the imaginary part of the Type I instability when Γ is decreased from 0.9 to 0.7, 0.6, 0.55 and 0.51, with which we see that the most unstable wavenumber moves to the long wavelength limit.

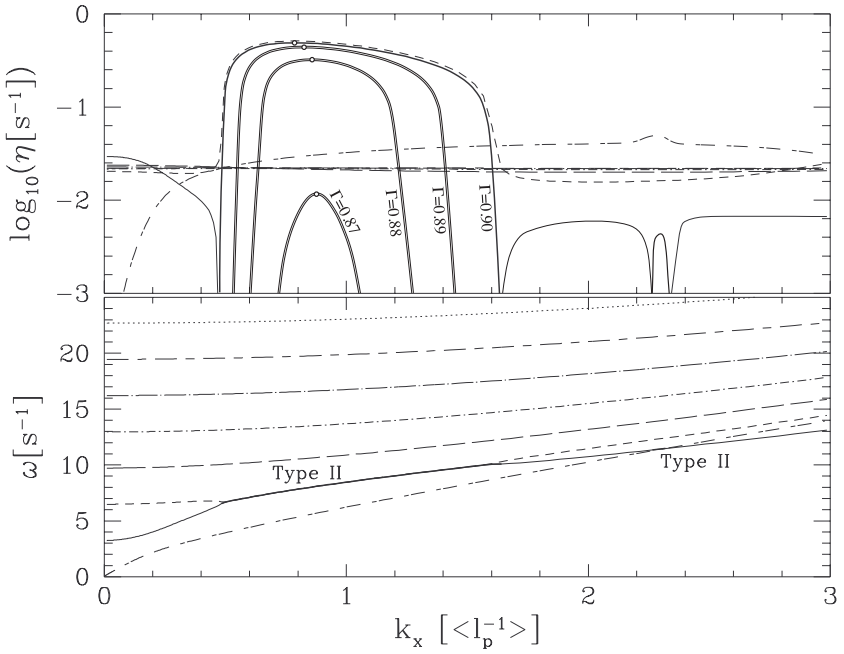

FIG. 5.— Same as the previous figure except that now the flux is kept constant as the bottom boundary condition. Here, the atmosphere is only unstable to the Type II instability. We see that the second eigen mode 'merges' at two different horizontal wavelengths, with either the third eigenmode or the first eigenmode. The additional lines represent the imaginary part of the growth rate as Γ is decreased. We see that the most unstable wavenumber changes only slightly with Γ. The growth rate is typically several times the oscillation period.

that is to say, there will be an anti-correlation between δE and $\delta\chi_v$. The coefficient of the anti-correlation isn't large. The largest anti-correlation is obtained if the lower boundary condition is fixed to be at $\tau_b = 2\sqrt{2}/3 \approx 0.94$ in which case $\delta E/E = -(3-2\sqrt{2})\delta\chi_v/\chi_{v,0} \approx 0.172\delta\chi_v/\chi_{v,0}$. The anti-correlation is larger than 0.1 for $5 \gtrsim \tau_b \gtrsim 0.2$.

A negative correlation implies that the radiation pressure is anti-correlated with the density.

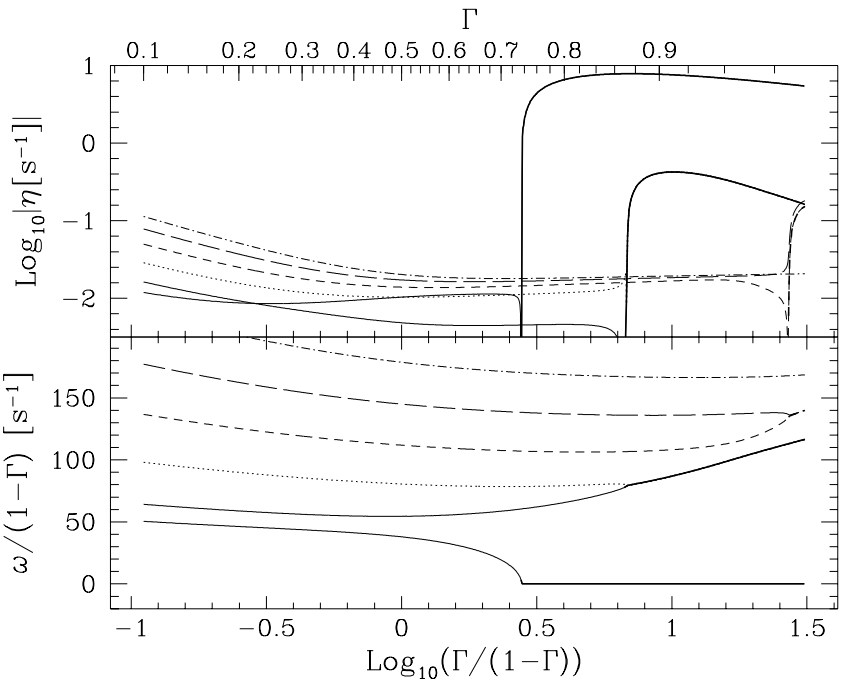

FIG. 6.— The eigenvalues of the first atmosphere described in Fig. 2 (with a fixed T at the bottom) as a function of Γ, while k_x is kept fixed at $\left\langle l_p^{-1} \right\rangle$. ω is multiplied by $1/(1-\Gamma)$ to normalize it to the changing dynamical time scale as the atmosphere puffs up when Γ is increased. The graph is actually plotted as a function of $\log_{10}(\Gamma/(1-\Gamma)) \approx \log_{10}\beta(\tau >> 1)$ in order to extend the high Γ region. Like in the previous figures, unstable modes are depicted with heavy lines. We see that as the flux is increased, this particular atmosphere first becomes unstable to the Type I instability (with a purely imaginary mode). As Γ approaches the Eddington limit, more and more modes become unstable to the Type II instability (which are mostly propagating).

When a fixed flux is imposed at the bottom, we have

$$F = F_0 = const$$
$$E = (2+3\tau)\frac{F_0}{c} \approx E_0(\tau_0)\left(1 + \frac{3\tau_0}{2+3\tau_0}\frac{\delta\chi_v(x)}{\chi_{v,0}}\right). \quad (51)$$

Here the correlation is positive, namely:

$$\frac{\delta E(\tau_0)}{E} \approx \frac{3\tau_0}{2+3\tau_0}\frac{\delta\chi_v(x)}{\chi_{v,0}}. \quad (52)$$

The coefficient of the correlation can be in this case larger than in the previous one and it increases with the optical thickness of the atmosphere. For $\tau_b \to \infty$, we have $\delta E/E = \delta\chi_v(x)/\chi_{v,0}$.

We conclude that under *the same local conditions* one can obtain qualitatively different behaviors according to the *boundary conditions* imposed. Should we have evaluated this atmosphere using a local analysis, we would have obtained that $\langle \delta p_{\rm rad} \rangle = 0$ since $k_z = 0$. Clearly, the boundary conditions are important to the question of stability.

In a real atmosphere with vertically dependent unperturbed state and modes, we could expect to see eigenmodes with both types of correlations. We point that the correlations do not depend on β and hence such phenomena can occur in low luminosity atmospheres. A typical example is the formation of grains in atmospheres. The appearance of grains causes large opacity changes.

5.2. *Toy model with non local effects*

We assume for simplicity that the relevant perturbations (of the order of the scale height of the atmosphere) are in the NAR limit. We also assume that the gas is held within a narrow slab without any z dependence. That is to say, we assume that the perturbation to the radiation density is proportional to the density perturbations:

$$E_1 = \varepsilon E_0 \frac{\delta\rho}{\rho} \quad \text{or equivalently} \quad p_{\rm rad,1} = \varepsilon p_{\rm rad,0}\frac{\delta\rho}{\rho}, \quad (53)$$

where ε is the 'efficiency' with which the average radiation energy density or pressure changes with changes in the opacity (which is a function of ρ). In §5.1, ϵ was calculated for a very simplified case. In reality, $\delta p_{\rm rad}$ is not simply proportional to $\delta\rho$ but is also a function of height, we just assume it to be so (this is why it is a 'Toy' model!). We therefore leave ε as an effective free parameter.

The equations describing the gas are the continuity and momentum equations. The linearized versions of which are

$$\frac{\partial\rho_1}{\partial t} = -\nabla\rho_0 \cdot \mathbf{v}_1 - (\nabla\cdot\mathbf{v}_1)\rho_0 \quad (54)$$
$$\rho_0\frac{\partial\mathbf{v}_1}{\partial t} = -\nabla p_1 - \nabla p_{\rm rad,1}. \quad (55)$$

Since the gas is in the NAR limit, its pressure perturbation is given by

$$p_1 = c_T^2\rho_1 + c_T^2\rho_0\frac{E_1}{4E_0} = \left(1+\frac{1}{4}\varepsilon\right)c_T^2\rho_1. \quad (56)$$

The radiation pressure perturbation can be written as

$$p_{\rm rad,1} = \beta p_0\frac{\rho_1}{\rho_0} = \beta c_T^2\rho_1. \quad (57)$$

Thus,

$$\frac{\partial^2\rho_1}{\partial t^2} - \Delta\rho_1\left(1+\varepsilon/4+\varepsilon\beta\right)c_T^2 = 0. \quad (58)$$

The first term in the parenthesis arises from the isothermal speed of sound and describes simple acoustic waves. The second term arises because the NAR limit is not necessarily the isothermal limit. Consequently, the residual correlation between δT and $\delta\rho$ imply that the relevant sound speed is modified to:

$$c^2_{\rm NAR} \equiv c^2_T(1+\varepsilon/4). \tag{59}$$

The third term within the parenthesis arises from the work that the radiation does in synchronization with the wave.

The corresponding dispersion relation is

$$\omega^2 = k^2\left(1+\varepsilon(\beta+1/4)\right)c^2_T \equiv k^2 c^2_{\rm eff}. \tag{60}$$

A negative ε implies a decrease in the effective speed of sound. When the anti-correlation is large enough, the speed of sound $c^2_{\rm eff}$ becomes imaginary and the waves simply amplify without propagating. In most astrophysical systems, this cannot arise unless β is sufficiently large because ε is seldom large by itself[1].

If ε is negative, the system will necessarily become unstable at some $L = L_{\rm crit} < L_{\rm Edd}$ for which β is large enough. Using the toy model and a typical value below the photosphere of $\varepsilon \approx -0.15$ which we should expect to see in an atmosphere with a τ_b of about 5, we have that the critical β in this case is $\beta_{\rm crit} \approx 1/\varepsilon - 1/4 \approx 6.5$, and the critical Eddington parameter is $\Gamma_{\rm crit} \approx \beta_{\rm crit}/(\beta_{\rm crit}+1) \approx 0.86$. The full numerical calculation gave $\Gamma_{\rm crit} = 1/2$ for a deeper atmosphere. This discrepancy can actually be reconciled exactly with a more elaborate toy model. The feature missing in this simple model is the 'feedback' that the change in the radiation flux has on the vertical extent of the atmosphere. Specifically, a lower density region will have a higher flux going through it and therefore a smaller effective gravity. Consequently, it would puff up. At a given height above the bottom, the radiation density will therefore be even higher because the Lagrangian coordinate will be lower. This will introduce a further anti-correlation between ρ and E. This missing feature in the toy model also allows very deep atmospheres to be unstable. Without it, we saw in §5.1 that as τ_b is increased, the anti-correlation tends to zero.

Once an atmosphere is unstable to Type I instabilities, it would tend to open 'chimneys' through which it is easier for the radiation to escape and also to accelerate material.

6. THE TYPE II INSTABILITY

To understand the Type II instability, it is best to compare it to the instability of radial s-modes (e.g., Glatzel 1994, and Papaloizou et al. 1997). As we shall see, there are many similarities and a few particular differences which will elucidate the origin of this instability.

The main similarities between s-modes and the Type I instability are:

1. The instabilities result with eigen-frequencies that describe modes with e-folding growth times of order the oscillation period. Each mode has a conjugate that is damped at the same rate. Moreover, as a control parameter is changed (e.g., $T_{\rm eff}$ in s-modes or Γ here), the conjugate pair arises when two modes with different real frequencies merge together.
2. Both types of instabilities appear only in systems in which the radiation pressure dominates.
3. Both instabilities are quite different from various 'slow instabilities' such as the κ-mechanism and other Carnot types.
4. Both instabilities disappear when the system is driven out of the two fluid limit. Namely, both need the radiation field to be described by the diffusion equations instead of being highly coupled to the gas.

Nevertheless, the are a few critical differences between the two:

1. The Type II instability operates in Thomson atmospheres, s-modes need a particular opacity law for them to be unstable. In particular, they need $\bar\kappa_T > 0$ which is obtained in ionization zones. Thomson atmospheres are much more general.
2. Unstable s-modes are a radial phenomenon. The Type II is intrinsically non-radial. Type II unstable modes do not exist if the horizontal wavelength is significantly larger than the pressure scale height of the system. It also implies that it will be a local phenomenon in stars since the scale height is usually much smaller than the radius.
3. s-modes are localized to the top part of the atmosphere, where the NAR limit is obtained, by having a region with an effective speed $c^2_\kappa < 0$. The Type II unstable modes will be localized to the top because they are a very high ℓ phenomenon.
4. Last, unstable s-modes require that the gas temperature be given by the radiation temperature, namely, the NAR limit. Type II unstable modes where found to exist also when the atmosphere was switched to the adiabatic limit (though still in the two fluid limit) by reducing the absorptive opacity enough to have $\tau_{\rm cool} \gg \tau_{\rm dyn}$.

We see that, mathematically, the origin of the two different instabilities is similar. The instabilities arise when the linearized equations become non self-adjoint (e.g., Papaloizou et al. 1997) as it is this characteristic that results in the complex conjugate pair of eigenmodes. In the case of s-modes, it was shown that it is because the dispersion equation becomes third instead of second order. Without the interaction with the diffusive radiation field, the pressure and density have a local algebraic relation of the form $p_1 = v_s^2\rho_1$. As such, the dispersion equation for ω^2 is self-adjoint and no complex roots are found for ω^2. However, because the temperature is set by the radiation field, the relation between density perturbation and total pressure perturbation becomes differential (Glatzel 1994). To see this, we look at the vertical component of eq. (5), and perturb it:

$$\frac{\partial p_{\rm tot,1}}{\partial z} \approx \frac{\partial p_{\rm rad,1}}{\partial z} = \frac{1}{3}\frac{\partial E_1}{\partial z} = -\frac{\chi_{v,1}F_{z,0}}{c} - \frac{\chi_{v,0}F_{z,1}}{c} \tag{61}$$

[1] However, the following are two known exceptions. (i) A photon + gas fluid with ν's playing the role of the radiation, (ii) An atmosphere with grains in which the opacity changes can be extremely large.

In the case of radial perturbations, $F_{z,1}$ vanishes. Thus, in s-modes, it is the first term on the r.h.s from which a differential relation is obtained between $p_{\rm tot,1}$ and ρ_1 which $\chi_{v,1}$ is a function of. The dispersion relation becomes third order and ω^2 can obtain complex conjugate roots. The differential nature of the pressure density relation is obtained by perturbing the radiation transport equation under a fixed flux, as it is in the two fluid limit (Glatzel 1994). Because the variable opacity, which is a function of the gas temperature, is one of the key ingredients needed to get the differential relation, it is clear why s-modes do not exist in the two fluid adiabatic limit. In this limit, the temperature of the gas, and therefore the opacity, is not a function of the radiation field but of the gas alone.

The physical origin of the Type II instability is similar to that of s-modes though not the same. The different characteristics of the two instabilities should point to the differences in physical origin.

First, like s-modes, the Type II doesn't exist in the one fluid limit and the radiation pressure must dominate, thus, it is probably the diffusive nature of the radiation field which promotes the Type II instability. Second, since Type II exists also in Thomson atmospheres, the form of the opacity is obviously not crucial. It also explains why the Type II exists also in the two fluid adiabatic limit – the instability is not sensitive to the properties of the gas, whether its temperature is set adiabatically or by the radiation. Thus, we hypothesize that the origin of the Type II instability is in the differential nature that the relation between the radiation pressure perturbation and the density perturbations obtains due to non-radial perturbations. The second term in eq. (61) fits the required features. It will not vanish if $F_{z,1} \neq 0$, namely, if there are perturbations to the flux in the vertical direction. Since $\nabla \cdot \mathbf{F} = 0$ in the diffusive limit, such a perturbation to $\mathbf{F}$ can arise only if there are horizontal perturbations $F_{x,1}$ as well. These arise only when there are horizontal perturbations to the density on the same scale as the vertical perturbations, hence the need for non-radial modes.

To check this hypothesis, we artificially modify in the radiation equations the term responsible for *horizontal* fluxes. Namely, we write $\nabla E \rightarrow (\partial E/\partial z)\hat{\mathbf{z}} + u_h(\partial E/\partial x)\hat{\mathbf{x}}$, where u_h is an artificial factor. For $u_h = 1$ we recover the correct diffusion equations. For $u_h \rightarrow 0$ we inhibit the flow of radiation in the horizontal direction. If our hypothesis is correct, then under the latter limit, the Type II instabilities should disappear. What we find in the numerical calculation is just that.

To summarize, the Type II instabilities originate from a differential relation between the total pressure and the gas density. This transforms the dispersion equations to a higher order equation in ω that can have complex conjugate roots. The differential 'equation of state' needed to relate the density and total pressure perturbations in s-modes, arises when perturbing the constant flux equations and it is sensitive to the opacity. The required differential relation in Type II unstable modes arises when horizontal perturbations are taken into account, allowing horizontal fluxes.

How does a Type II unstable system look like? To see this, we plot in Fig. 7 a snapshot of the eigenmode of the system. Plotted is the energy density of the radiation as a function of x or $-t$. The solution is proportional to $E_1(z)\exp(ik_x x - i\Re(\omega)t)\exp(\Im(\omega)t)$. Thus, if we fix t, we obtain E_1 which varies with z and is harmonic in the x direction. If we vary the snapshot with time, we will find that it propagates to the right side. If we fix x, let t and z vary, and exclude the exponential growth, we will obtain the same figure, with the horizontal axis for the time flipped. Namely, we will see at a fixed x, structure moving *downwards*. This is opposite to conventional photon bubble picture.

7. DISCUSSION

We have seen that atmospheres with a significant radiative support always become unstable if the radiative pressure is increased sufficiently and becomes dominant. Both instabilities found are intrinsically non radial. Table 1 summarizes some of the characteristics of the two instabilities found and emphasizes the differences between them and other known instabilities – convection and s-modes which are both dynamic instabilities, and the κ-mechanism which is a Carnot type instability.

Using a Toy model, we can understand the Type I instability. The anti-correlation between gas density and total pressure, dominated by the radiation, reduces the effective speed of sound. If the reduction is large enough, the speed of sound squared becomes negative and we have a purely imaginary frequency. This explains why $\Im(\omega)$ was found to be proportional to k_x at long wavelengths.

We have also seen with the Toy model that in addition to the radiative pressure term, changes to the temperature can contribute towards an instability. In Thomson atmospheres, this term is never large enough. However, when opacity variations can be very large, as is the case with dust formation, the temperature perturbation can by itself make the speed of sound become imaginary. In such a case, the lower limit for an instability is the lowest flux for which oscillations on the dynamical time scale are still not adiabatic.

The 'temperature' term contribution to the effective speed of sound can become important in a second type of cases in which the equation of state becomes significantly 'softer' than that of an ideal gas. That is, when $\partial \ln p/\partial \ln T|_\rho \gg \partial \ln p/\partial \ln \rho|_T$. For example, in a system in which the fluid is composed of highly coupled radiation and plasma and the role of the radiation is played by a strong ν flux, then given temperature variations will now be more important in setting the effective speed of sound. When this happens, we can again expect to see an instability even in systems which are far from the Eddington flux (as long as oscillations on dynamical time scales are not adiabatic).

The second type of instabilities found, the Type II, were found to be similar to s-modes. The difference however is that Type II does not require special opacity laws on one hand, but on the other hand, it does require non radial oscillations. The latter allows for a differential relation between the radiation pressure perturbations and the gas density which can result with a non self-adjoint dispersion equation for ω^2. Once triggered, the unstable modes describe highly distorted horizontally propagating waves as is depicted in Fig. 7. Interestingly, if a standing

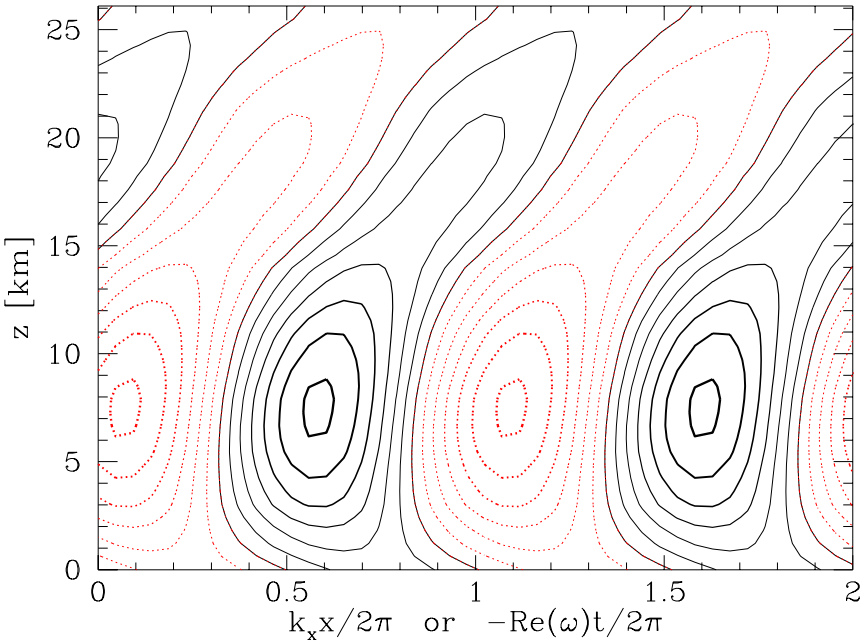

FIG. 7.— The radiation energy density of the Type II unstable mode of the 2nd atmosphere of Fig. 2 as a function of z and either x or $-t$. If we choose the horizontal axis to be x, we see a snapshot of the eigenmode. Solid lines are positive iso-contours of the radiation energy density while dotted are negative. The normalization is meaningless in the linear problem. With time, this pattern moves to the right and amplifies according to $exp(\eta t)$. If we look at the evolution of the profile along a vertical cut, its evolution of $\tilde{E}(z,t)$ without the exponential growth will be given by the same plot with the horizontal axis being $-t$, namely, it would appear as if structure is moving *downwards*.

TABLE 1
COMPARISON BETWEEN THE INSTABILITIES FOUND HERE AND THREE OTHERS.

Instability:	Type I	Type II	Convection	s-mode	κ-mechanism
Radial (exists for $k_x \to 0$)	-	-	-	+	+
Propagating ($\Re(\omega) \neq 0$):	-	+	-	+	+
Dynamic time scales ($\Im(\omega) \sim \mathcal{O}(1/\tau_{\rm dyn})$)	+	+	+	+	-
Exists in adiabatic limit ($\tau_{\rm cool}/\tau_{\rm dyn} \to \infty$)	+	+	+	-	[a]
Exists in NAR limit ($\tau_{\rm cool}/\tau_{\rm dyn} \to 0$)	+	+	-	+	[a]
Exists in diffusion limit ($\tau_{\rm diff}/\tau_{\rm dyn} \to 0$)	+	+	+	+	[a]
Exists in one fluid limit ($\tau_{\rm diff}/\tau_{\rm dyn} \to \infty$)	-	-	+	-	[a]
Requires large $p_{\rm rad}/p_{\rm gas}$	+	+	-	+	-
Requires special opacity laws	-	-	-	+	+
Significantly affected by boundaries	+	-	-	-	-

[a] The κ-mechanism per se requires that the gas will *not* be fully in the adiabatic limit. Other similar Carnot type instabilities can exist when the gas is *not* fully in the other three limits. Namely, Carnot type instabilities can potentially exist if the gas is not in one of the four corners of fig. 1.

wave in the horizontal direction is composed out of two oppositely propagating waves, the result appears as 'bubbles' propagating *downwards*! That is, it appears as 'anti-bubbles'. The common conception that a 'bubbly' phenomenon should arise as the Eddington limit is approached assumes that the 'vacated' bubble through which it is easier for the radiation to escape can be held together by something (Spiegel 1977). This is true in the case of 'photon bubbles' in highly magnetized media where the material is forced to move only radially by vertical magnetic fields (Arons 1992). When no magnetic field is present, nothing is found to hold the material (Spiegel 1977), and as a result, these bubbles probably do not exit at all. What we find instead is a very dynamic picture in which the structure or 'phase speed' is actually downwards, such that the material in the 'vacated' regions is constantly changing on dynamical time scales. This is the opposite of bubbles. Of course, this pattern movement downwards does not involve net motion of material, and it facilitates the transfer of radiation upwards. What the exact form that this 'anti-bubbles' will look like, when in the nonlinear regime, is of course an interesting open question that requires a full numerical simulation.

A very important question which should be addressed in future work is the effects that the assumption of trapping of modes has. That is, it was assumed in the form of the boundary conditions, that acoustic waves are reflected at the boundaries above and below the layer. This need not

necessarily be the case. If the atmosphere does not have a corona for example, where the temperature and speed of sound increase with height, any sound wave traveling through the atmosphere will just continue upwards and dissipate itself after forming shocks (which could in principle generate the required corona). In any case, we expect both instabilities to be important because they amplify on a short dynamical time scale – a mode does not need to travel back and forth to be amplified significantly. Moreover, the imaginary effective speed of sound in the Type I instability will be shown in a subsequent publication to be a signature of a phase transition in which the homogeneous state of the atmosphere is no longer the preferred or lowest energy state, that is, the instability found is just a manifestation of the system trying to get into a new equilibrium condition which is not the homogeneous one.

Although a lengthy description of the nonlinear behavior is part of a forthcoming publication, a few words are in order about how such systems are expected to behave once these instabilities start to set in.

The results of §4.3.3 have shown that either the Type I or the Type II instabilities will necessarily set in before the Eddington limit is reached. Thus, as a luminous object approaches the Eddington limit, convection will become important in the inner parts of the object where the density is high enough and the cooling time scale long (Joss, Salpeter & Ostriker 1973), while the outer parts will be unstable to the Type I or II instabilities where the two fluid limit is relevant, even if the opacity is Thomson dominated. Once either one of the instabilities takes place, it is clear that the atmosphere will not remain homogeneous. If the nonlinear pattern that is eventually formed propagates, as is the case in the Type II instability, each gas element will experience periodic density variations. While the peak force experienced by each mass element may exceed the Eddington limit, the time average force may be significantly below it (Shaviv 1998). In this case, an atmosphere can transfer through it a luminosity greater than the Eddington luminosity without blowing itself apart. This will occur in the region of the atmosphere where perturbations of order the scale height become optically thin. From this region upwards the effective opacity is again the microscopic one and the luminosity becomes effectively super Eddington. A predictable wind will then be blown from that region upwards (Shaviv 2000b).

In the case of the Type I instability, on the other hand, the pattern formed is stationary and it resembles that of 'chimneys'. The gas elements present in the lower density regions can witness a constant (in time) super-Eddington flux. Although the average over the entire atmosphere can be a sub-Eddington luminosity, some gas elements in this case can be accelerated upwards through the chimneys, driving turbulence in the shear layers along the sides of the chimneys. However, this material will probably not be able to escape to infinity because it is unlikely that the escape speed will be attainable in the chimneys. This will require strong shears that presumable dissipation by shocks will prevent it from occuring. Thus, once the Eddington flux is surpassed in the chimneys, it will likely generate 'fountains' in which material accelerates upwards until the top of the chimneys is reached, then the material will fall back. If the average luminosity is super Eddington, the material need not stagnate and it can be easily driven off to infinity as a wind.

In both cases, the clear role that the classical Eddington limit plays in the homogeneous case is lost. On one hand, a strong wind can exist already below the Eddington limit, driven for example b the energy dissipated from the unstable modes; while on the other hand, quasi-stationary configurations (with a wind) can exist above the Eddington limit.

A more detailed treatment of the nonlinear behavior, which includes toy models as well as numeral simulations is underway in a forthcoming publication.

8. SUMMARY

The two main results found in this paper are the existence of two types of dynamically important instabilities that can take place in Thomson atmospheres. The two instabilities exist in addition to the possibility of convection. Following is a summary of the main results pertaining to these two important instabilities which will govern the behavior of atmospheres shining close to the Eddington limit.

1. Under a linear analysis of a slab Thomson atmosphere, two dynamically important instabilities were found to exist for the first time.

2. Both instabilities operate in the two fluid limit of optically thick atmospheres – when the diffusive time scale is shorter than the dynamical time scale of the system. Hence, they operate from the photosphere down to where convection can become efficient.

3. The Type I instability is discovered once a nonlocal analysis of the radiation field behavior is carried out. A fixed flux boundary condition quenches it off.

4. The Type I instability does not propagate. It corresponds to the opening of 'chimneys'. It appears in modes that have the radiation density on average anti-correlated with the gas density. The anti-correlation reduces the square of the effective speed of sound. When $c^2_{\rm eff}$ is negative, the instability appears and grows over a dynamical time scale.

5. The Type I instability can take place also for $k_x \to 0$, however, the growth rate then tends to vanish: $\eta \propto k_x \to 0$.

6. A second instability, the Type II, appears under more general boundary conditions, but at somewhat higher Γ's than the Type I.

7. The mathematical origin of the Type II instability is similar to that of s-modes. Since the hydrodynamic equations are coupled to a diffusion equation for the radiation, the relation between the pressure and density is differential. In the Type II instability, it arises from the term responsible for generating perpendicular radiative fluxes.

8. The Type II instability describes horizontally propagating waves that amplify over dynamical time scale. If a standing mode is composed out of two oppositely propagating modes, it has the appearance of 'bubbles'

moving downwards. This arises because the phase velocity of the wave standing in both the horizontal and vertical direction is pointing down. This is opposite to the intuitive 'photon bubble' picture, in which the structure appears to be moving upwards.

9. As the flux through an atmosphere increases towards the Eddington limit, Thomson scattering dominated atmospheres become unstable above a critical fraction of the Eddington flux. This instability is independent of and comes on top of the possibility of convective instability. The first mode to become unstable can be either Type I or Type II unstable. Type I appears at $\Gamma \gtrsim 0.5$. However, the exact value depends on the boundary conditions and thermal gradient. The Type II instability appears under more general boundary conditions but at $\Gamma \approx 0.85$.

10. The most unstable horizontal wavelength for the type I instability is the longest wavelength in the system at $\Gamma_{\rm crit}$, but it becomes smaller as Γ is increased. The most unstable mode for the type II instability is about $0.9\left\langle l_p^{-1}\right\rangle$. It doesn't change by much as Γ is increased from $\Gamma_{\rm crit}$.

11. The analysis assumed thus far that the modes can be trapped around the layer under consideration. However, since the two instabilities operate over a dynamical time scale, they are probably important even if the acoustic oscillations are not trapped, as is the case with convection for example.

12. The unstable modes always tend to reduce the effective opacity of the medium (if the luminosity is sufficiently close to the Eddington luminosity), thereby increasing the effective Eddington luminosity. The radiation and the opacity must therefore be solved simultaneously and self consistently. Treatment of the inhomogeneities on a large scale (say scales larger than a few optical lengths) requires a new formulation of the equation of state because of the process of averaging over the inhomogeneities. This problem is deferred to a later paper.

All the results presented in this work assume that the mass opacity is constant. In general however, the opacity can be a function of the gas variables. This can be shown to not only change quantitatively the results but also to introduce the effects of additional instabilities, some of which have not yet been described. Moreover, it should be stressed that since the temperature of the matter and the radiation at any given point are not necessarily identical, non equilibrium effects cannot be neglected. These effects should be discussed in the future.

Finally, if we are also to fully understand the behavior of the instabilities, we should analyze their nonlinear behavior. It is this non linear behavior that will set many of the observational properties of systems shining close to their Eddington limit.

The author wishes to thank the comments and suggestions made by the editor, anonymous referee, Mitch Begelman and Roger Blandford as well as Caltech for the DuBridge Prize fellowship which supported him and CITA for the current support.

APPENDIX: A THIRD INSTABILITY?

The numerical linear analysis performed by Spiegel & Tao (1999) (hereafter S&T) is similar to the analysis carried out here. However, S&T report on one hand the finding of a third instability that is not found here. On the other hand, they do not find the two instabilities that are described in this work. This of course raises the following two questions. Why didn't they find the instabilities described here and why didn't the analysis described here result with their instability. We wish to address these two questions here.

The answer to the first question is rather simple. S&T only analyze one type of boundary conditions. The particular boundary conditions they chose tend to promote a lowest eigenmode with a net positive correlation between the radiation pressure and the gas density. This can be seen in their eigenmode profiles by comparing the radiation energy density and the gas pressure. Thus, their boundary conditions do not allow the 'Type I' instability to show up.

The second instability, the 'Type II', was found to operate on higher vertical eigenmodes than the lowest one. Since S&T used a relaxation method to solve the eigenvalue problem and not a shooting method as was carried out here, they were limited to the lowest eigenmode only. Thus, they couldn't find the second instability either.

The instability found by S&T is similar to various 'overstability' instabilities, such as the κ-mechanism, in which the growth time scale is significantly longer than the oscillation period. In their case, they found growth rates of order 100's to 1000's of oscillation periods. Why wasn't this instability found here?

We searched with our code the available parameter space and no instability, as the one described, was found. One should note that S&T themselves cite the unpublished work of Marzek (1977) who studied the same problem but could not find this instability either. We could, however, artificially get an instability similar to the one described in S&T, if we did one of the following:

1. Calculate the Doppler term wrong, or neglect it altogether.

2. Use an inaccurate unperturbed atmosphere.

3. Have a mismatched opacity between the unperturbed atmosphere and the perturbations.

Since S&T found in their analysis that the Doppler terms result with only a very minor correction to the eigenmodes, the analysis described in this work was incorrectly first executed without the Doppler terms. It was then found that the finite speed of light terms – those multiplied by u_c, result with an instability very similar to what was described by S&T. When we added the Doppler terms, we found that the radiation drag that it induces is always comparable but larger than the instability arising from the u_c terms. An analytical estimate for the size of the terms corroborates this and it shows that the absorption term, which introduces damping as well, can be either larger or smaller. Namely, there are two possibilities. In the first,

S&T calculated the Doppler terms correctly and the origin of the instability in not with the u_c terms, in which case they must have used a very small v_s/c ratio (found in a low temperature atmosphere) a fact which is not stated. The second possibility is that the instability does originate with the u_c terms in which case they erroneously underestimated the Doppler term. One should note that S&T use the Eulerian description of the radiation while we use the Lagrangian description, namely, we write the zeroth and first moments of the radiation field in the frame of reference of the moving material. The former method is, as it turns out, significantly more complicated since it involves many terms instead of one.

If the Doppler terms were calculated properly, a second way of artificially getting an instability is if the atmosphere used as the zeroth order, or unperturbed state, is inaccurate and doesn't solve the radiation equations accurately. S&T describe the exact unperturbed solution, which for the optically thin limits is isothermal and the thick limits polytropic. They then solve for the eigenmodes assuming one or the other and then in an atmosphere that has both regions. If the latter case is actually not the accurate solution but the two limits 'glued' together, it would explain how an instability can arise. This is because the unperturbed state does not solve exactly the unperturbed equations. The modes display an instability 'triggered' by the atmosphere's will to satisfy the hydrostatic equations. We have also found the same effect for the same reasons if the unperturbed atmosphere is calculated with only the scattering opacity (which dominates), while the linear analysis is calculated with the scattering and very small absorption opacity.

Irrespective of what the resolution to this question is, even if the instability does really exist, it is dynamically less important than both the Type I and Type II instabilities.

REFERENCES

Appenzeller, I. 1986, in Luminous Stars and Associations in Galaxies, IAU Symposium No. 116, ed. C. W. H. de Loore, A. J. Willis, and P. Laskerides (Dordrecht, Reidel), p. 139
Asplund, M. 1998, A&A, 330, 641
Arons, J. 1992, ApJ, 388, 561
Berthomieu G., Provost J., & Rocca A. 1976, A&A, 47, 413
Carlberg R. G. 1980, ApJ, 241, 1131
Eddington, A. S. 1926, The internal constitution of stars, Cambridge Univ. Press, Cambridge
Feldmeier A., Puls, J., & Pauldrach, A. W. A. 1997, A&A, 332, 878
Gammie, C. F. 1992, MNRAS, 297, 929
Gautschy, A.,& Glatzel, W. 1990, MNRAS, 245, 597
Glatzel, W., & Kiriakidis, M. 1993, MNRAS, 263, 375
Glatzel, W. 1994, MNRAS, 271, 66
Hearn, A. G. 1972, A&A, 19, 417
Hearn, A. G. 1973, A&A, 23, 97
Humphreys, R. M., & Davidson, K. 1984, Science, 223, 243
Humphreys, R. M., & Davidson, K. 1994, PASP, 106, 1025
Joss, P.C., Salpeter, E.E. and Ostriker, J.P. 1973, Ap.J. 181, 429
Landau, L. D. & Lifshitz, E. M. 1987, Fluid Mechanics 2nd ed., Pergamon Press, Oxford.
Lamers, H. J. G. L. M 1986, in Luminous Stars and Associations in Galaxies, IAU Symposium No. 116, ed. C. W. H. de Loore, A. J. Willis, and P. Laskerides (Dordrecht, Reidel), p. 157
Ledoux, P. 1965, in Luminous Stars, L. H. Allter and D. B. McLaughlin, The university of Chicago Press, Chicago
Marzek, C. 1977, Thesis, Dept. of Astronomy, Columbia University.
Massey, P. & Hunter 1998, D. ApJ, 493, 180
Mihalas D., & Weibel Mihalas B. 1984, Foundations of radiation hydrodynamics, Oxford univ. press, Oxford
Owocki, S. P. & Rybicki, G. B. 1984, ApJ, 284, 337
Owocki, S. P. & Rybicki, G. B. 1985, ApJ, 299, 265
Owocki, S. P. & Rybicki, G. B. 1986, ApJ, 309, 127
Owocki, S. P. & Rybicki, G. B. 1991, ApJ, 368, 261
Papaloizou, J. C. B., Alberts, F., Pringle, J. E., & Savonije, G. J. 1997, MNRAS, 284, 821
Prendergast, K. H., & Spiegel, E. A. 1973, Comm. Ap. Space. Phys., 5, 43
Rybicki, G. B., Owocki, S. P. & Castor, J. I. 1990, ApJ, 349, 274
Saio, H., Baker, N. H., Gautschy, A. 1998, MNRAS, 294, 622
Shara, M. M., Zurek, D. R., Williams, R. E., Prialnik, D., Gilmozzi, R., Moffat, A. F. J. 1997, AJ, 114, 258
Shaviv, N. J. 1998, ApJ, 494, 193
Shaviv, N. J. 1999, in Variable and Non-spherical Stellar Winds in Luminous Hot Stars, IAU Colloquium 169, ed. B. Wolf, O. Stahl, A. W. Fullerton, Springer, p. 155
Shaviv, N. J. 1999, Phys. Rep., 311, 177
Shaviv, N. J. 2000, ApJ, 532, L137
Shaviv, N. J. 2000, submitted to MNRAS
Spiegel, E. A. 1976, in: Physique des mouvements dans les atmospherès stellaires, ed. Cayrel R. & Steinberg M., CNRS, Paris, p. 19
Spiegel, E. A. 1977, in: Problems in Stellar Convection, eds. Spiegel, E. A., & Zahn, J.-P., Springer-Verlag, Berlin.
Spiegel, E. A., & Tao L., 1999, Phys. Rep., 311, 163.
Stothers, R. B., & Chin C.-W. 1993, ApJ, 408, L85
van Genderen, A. M., & Thé, P. S. 1985, Space Sci. Rev , 391, 317
Wagenhuber J., & Weiss A. 1994, A&A, 290, 807